PATENT
Docket No.: 19603/461 (CRF D-1595A)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Barany et al.	Examiner: J.W. Ricigliano
Serial No.	:	08/794,851	Art Unit: 1648
Filed	:	February 4, 1997	
For	:	DETECTION OF NUCLEIC ACID SEQUENCE DIFFERENCES USING THE LIGASE DETECTION REACTION WITH ADDRESSABLE ARRAYS	

DECLARATION OF FRANCIS BARANY UNDER 37 CFR § 1.132

Assistant Commissioner for Patents
Washington, D.C. 20231

Dear Sir:

I, FRANCIS BARANY, declare:

1. I received a B.A. degree in Chemistry from the University of Illinois at Chicago Circle, Illinois in 1976 and a Ph.D. degree in Microbiology from the Rockefeller University, New York, New York in 1981. I conducted postdoctoral work from 1981 to 1982 in microbiology at the Rockefeller University, New York, New York and from 1982 to 1985 in molecular biology at the Johns Hopkins University School of Medicine, Baltimore, Maryland.

2. I am a Professor, Department of Microbiology, Cornell University Medical College, New York, New York and an Adjunct Professor at the Rockefeller University, New York, New York.

3. I am a named inventor of the above-identified patent application.

4. I submit this declaration to demonstrate that it would not have been obvious to arrive at the invention of my above-identified application by combining prior art ligase detection reaction technology with nucleotide array technology.

5. The ligase chain reaction ("LCR") procedure is useful in detecting single-base differences in target nucleotides. LCR involves the use of two sets f oligonucleotide probes to achieve exponential amplification. The LCR procedure is distinguishable from the ligase detection reaction ("LDR") technique by virtue of the fact that LDR only uses a single set of oligonucleotide probes which will hybridize to only one target nucleic acid strand and, thus, achieve linear amplification. These procedures are discussed e.g., in Barany, "Genetic Disease Detection and DNA Amplification Using Cloned Thermostable Ligase," Proc. Natl. Acad. Sci. USA, 88:189-193 (1991), Barany, "The Ligase Chain Reaction in a PCR World," PCR Methods and Applications, pp. 5-16 (1991), and Wiedmann, et al., "Ligase Chain Reaction (LCR) – Overview and Applications," PCR Methods and Applications, S51-S64 (1994) ("Wiedmann"), all of which are of record in my present application.

6. LDR and LCR are useful in detecting genetic diseases where a genetic defect is either not present (i.e., normal individuals), present in 50% of the DNA (i.e., heterozygous individuals), or is always present (i.e., homozygous mutants). However, when dealing with cancer and cancer-associated mutations, the detection problem becomes more complex. In primary cancer tumors, signal may be as low as 15% of wild-type DNA. For early detection of cancer or detection of metastasis, signal may be in the range of 1% to 0.1% of wild-type DNA.

7. Cancers arise from the accumulation of multiple mutations in cellular growth- and differentiation-regulation genes. Oncogenes may be activated by point mutations, translocation, or gene amplification, while tumor suppressor genes may be inactivated by point mutations, frameshift mutations, and deletions. These mutations may be inherited or somatic, arising from exposure to environmental factors, or arising from a malfunction in DNA replication and repair machinery. There is an urgent need to apply new DNA tools to further our understanding of the development and progression of cancers.

8. Identifying mutations responsible for cancer development presents multiple challenges. Many genes have been associated with cancer development, and no single gene is mutated in all cases of a particular adult cancer. As an example, the three most commonly mutated genes in colorectal cancer are the APC, p53, and K-*ras* genes, found mutated in approximately 70%, 50%, and 40% of colorectal cancers. Few individuals have mutations in all three genes, and some do not have a mutation in any of these genes. For

these three genes, there is some clustering of mutations i.e. "hot spots," but a mutation detection scheme needs to detect multiple potential mutations in multiple codons in multiple genes. See Figure 1 below.

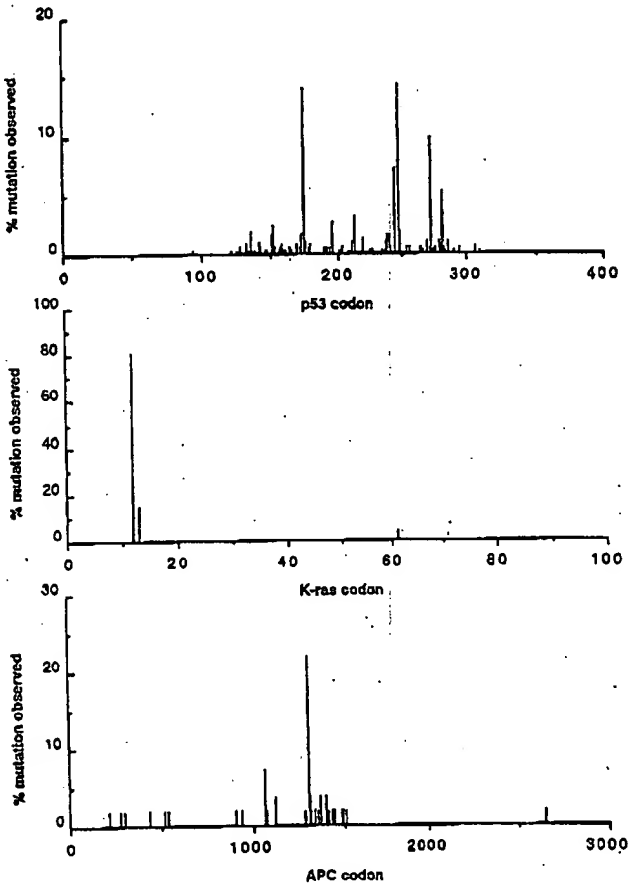


Fig. 1. Histogram of mutations observed in the p53, K-ras and APC genes in colon cancer.

Further, in primary tumors, normal stromal cell contamination can be as high as 70% of total cells, and thus a mutation present in only one of the two chromosomes of a tumor cell may represent only 15% of the DNA sequence present in a sample for that gene. This frequency is too low for reproducible detection by either standard automated fluorescent dideoxysequencing reactions or allele specific oligonucleotide hybridization ("ASO").

Finally, for true early detection of mutations, a technique needs to be developed to find such mutations in peritoneal or colonic lavage specimens, stool, or other bodily fluids. Preliminary work on stool samples shows that such mutations are present at about the 1% level, far too low for detection by either fluorescent dideoxysequencing or allele specific oligonucleotide hybridization. Thus, there is an urgent need to develop technology which can accurately quantify one or more low abundance mutations, at multiple adjacent, nearby, and distal loci in multiple genes.

9. A number of array-base technologies are known in the art. As demonstrated below, these techniques are not satisfactory for cancer detection. In U.S. Patent No. 5,744,305 to Fodor ("Fodor"), which is of record, and Pease "Light-Generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis," Proc. Natl. Acad. Sci. USA, 91:5022-26 (1994) ("Pease") (attached hereto as Exhibit 1), strategies to manufacture arrays are described. The simplest use of such arrays is for quantifying expression levels of different genes. In a simple method, which relies on fragment capture, a previously-labeled fragment containing a selected sequence is captured by hybridization to an immobilized probe. See Figures 2A-B below.

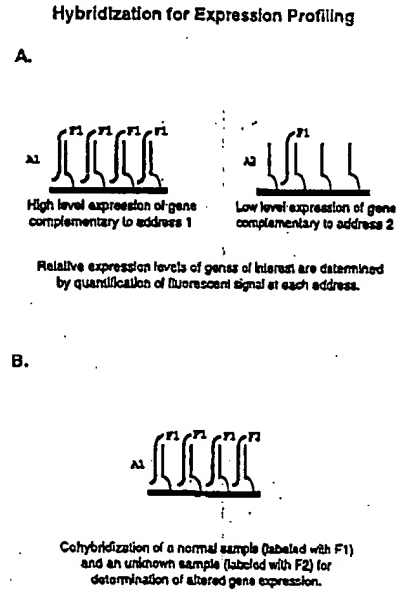


Fig. 2. DNA array to detect gene expression levels.

Variations of this technique have been developed in an attempt to quantify the precise amount of each gene. In the first approach, multiple oligonucleotides which represent sequential 20mer complements of each gene are placed on an array (Figure 2A). Although different sequences hybridize with different intensities, by comparing the relative intensities of multiple regions from two genes in one sample with their counterparts in a control sample, one can determine the relative expression levels of the genes. In a second approach, genes from one sample are labeled with a first fluorescent group and genes from a second sample are labeled with a second fluorescent group, and the two samples co-hybridize on an array containing cDNA complements of about 300 bases for each gene (Figure 2B). The relative intensities of the two fluorescent groups at each address are used to detect changes in gene expression levels. While these hybridization techniques are useful for determining presence or absence of a given nucleotide sequence target, and roughly quantifying the relative expression level compared with other nucleotide sequence targets of substantially different sequences, it is not useful for distinguishing single nucleotide differences.

10. A second use of the Fodor array technology is to determine nucleotide sequences of short fragments. This technique, known as "sequencing by hybridization" aims to place the complete set of 65,536 possible octameric oligonucleotides on a glass surface, and then determine the sequence of a short fragment by assembly of sequences based on which addresses the labeled fragment hybridizes to. A synthetic strategy to manufacture an array with 65,536 addresses is provided in Fodor '305 and Pease. The technique has not been successfully reduced to practice, because fragments hybridize to incorrect addresses (with one or more mismatches) in an unpredictable manner. An alternative idea is to tile across a region of a gene which may undergo deletion with nonamer (9-mer) sequences, such as exons 70 and 74 of the Duchenne Muscular Dystrophy gene (Beattie et al., "Advances in Genosensor Research," Clinical Chemistry, 41:700-706 (1994) ("Beattie") (attached hereto as Exhibit 2)). Exon specific probes hybridize to both correct (same exon) and incorrect (other exon) addresses. The authors note that specificity of hybridization could be improved somewhat by rendering the target single-stranded using heat denaturation as follows:

> Heat denaturation of the target strands had two beneficial effects: increasing the number of specific (expected) hybridization signals, and decreasing the number of nonspecific (unexpected) hybridization signals.

Id. at 703. Figures 3A-B show this result, with wide variation in signal intensity and several incorrect hybridizations.

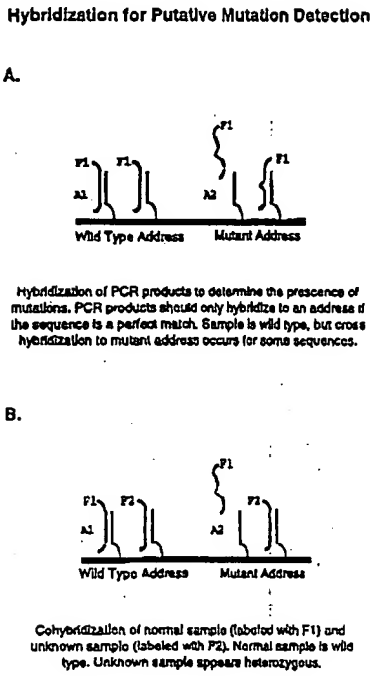


Fig. 3. DNA array using hybridization to detect putative mutation.

Thus, even when the target sequence is known, "sequencing by hybridization" does not determine the correct sequence. Guo et al., "Direct Fluorescence Analysis of Genetic Polymorphisms by Hybridization With Oligonucleotide Arrays on Glass Supports," Nucl. Acids Res., 22:5456-5465 (1994) ("Guo"), which is of record, attempted to overcome some of these technical difficulties by testing numerous conditions to detect known point mutations in the human tyrosinase gene. Target DNA was PCR amplified with one primer fluorescently labeled, and the other labeled with biotin. They required an extra step to separate the two strands of the PCR product. They tested a number of different linker lengths and probe lengths, and concluded that poly dT spacers of 15 or longer were required, and optimal hybridizations were obtained with probe addresses of 15 nucleotides with a GC content of 5-7 nucleotides. While these specialized conditions worked for the limited number of mutations tested, subsequent workers testing a larger range of sequences under a single

hybridization step have not obtained the same success rate (see below). Further, mutations in cancer genes often occur in GC rich regions, such as the K-*ras* codons 12 and 13 mutations frequently mutated in colon cancers. Thus, Guo is not generally applicable for determining the presence of all types of mutations on a single array.

11. A more complete test to determine sequence variation by hybridization uses a set of addresses for a given gene containing overlapping sequences fifteen bases in length, each address differing from the previous one by moving a single base pair along the gene sequence, i.e. Address 1 = nt_{1-15}; Address 2 = nt_{2-16}; Address 3 = nt_{3-17}; etc. Each address is duplicated four times, with the central base varied to give the wild-type base and the three possible mutations at that position. In theory, a labeled target will hybridize preferentially to the perfectly matched sequence at each address. The central base from each address will interrogate the target and will in theory determine the presence or absence of a mutation. See Figure 3A. Such an approach was used to attempt to tile through the mutations in exon 11 of the Cystic Fibrosis gene (Cronin et al., "Cystic Fibrosis Mutation Detection by Hybridization to Light-Generated DNA Probe Arrays," Human Mutation, 7:244-255 (1996) ("Cronin") (attached hereto as Exhibit 3)). While the wild-type sequence hybridize as predicted, Cronin writes (pg. 248):

> [H]owever, it is also evident that signal intensities of perfect match hybridizations vary, as do mismatch probe hybridization intensities. This is because different probes in the array have different thermal stabilities. Absolute signal intensities within a hybridized array vary with factors such as relative thermal stability of each duplex, hybridization conditions, and target concentration.

In other words, while some parts of the array hybridized very well, other parts hybridized very poorly. In addition, hybridization signal was observed at mutant sequences when only wild-type DNA was present (See Figure 4A).

12. Cronin writes (pg. 248):

> Considering these results, it is clear that a single base alteration in the context of a long DNA target sequence may sometimes be difficult to distinguish unequivocally by using single sets of interrogation probes to check every nucleotide position in a target sequence. This is particularly true with heterozygous genomic samples, where wild-type and mutant sequences are

present in equal amounts and hybridize with similar intensity to their array complements.

In other words, this array does not always work for detecting known mutations in heterozygous samples, where mutant and wild-type DNA are present in equal 50% quantities.

13. A very complex set of 40 addresses may be constructed to distinguish known germline mutations from wild-type (See Cronin); however this approach is still not amenable for detecting a minority signal. An alternative approach to using 40 addresses for each known mutation is to tile across the sequence (such as the very large BRCA1 gene) and label normal DNA with a first fluorescent group and the unknown DNA with a second fluorescent group and compare signal intensities across each position (Figure 3B). While this approach was somewhat better in detecting frameshift mutations, it is of insufficient sensitivity for determining the presence of mutations in primary tumors. One of the major drawbacks of this technique is the requirement to fragment the labeled target to achieve hybridization to the oligonucleotide addresses on the surface. This causes some inappropriate signals from partial hybridizations of smaller fragments to incorrect addresses. If the probe is left intact, differences in hybridizations to tiled 17mer sequences may vary 5-fold with no obvious features in the sequence to predict this variation in hybridization (Milner et al., "Selecting Effective Antisense Reagents on Combinatorial Oligonucleotide Arrays," Nature Biotechnology, 15:537-541 (1997) ("Milner") (attached hereto as Exhibit 4)) Finally, a recent article (Wang et al., "Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome," Science, 280:1077-1082 (1998) ("Wang") (attached hereto as Exhibit 5)) used the hybridization array described above for tiling through a sequence (Figure 3A). Initial hybridization to thousands of single-nucleotide polymorphisms (i.e. single-base mutations) gave a false-positive rate of 12% and false-negative rate of 13%, see id. at 1079, even though a given polymorphism is present in at least 50% of the target DNA. Such a high false-positive rate would be of questionable clinical utility for determining cancer mutations. In summary, this hybridization technique will not be able to unequivocally distinguish the presence of cancer mutations in primary tumors (where the mutation may represent 15% compared to 85% wild-type sequence), and certainly will not work for clinical samples, where the mutation represents only 1% or less of the target sequence.

14. Seeking to vercome the limitations of detecting single-nucleotide sequence differences by hybridization alone, some researchers have attempted to use a solid-phase primer extension approach (See Figure 4A).

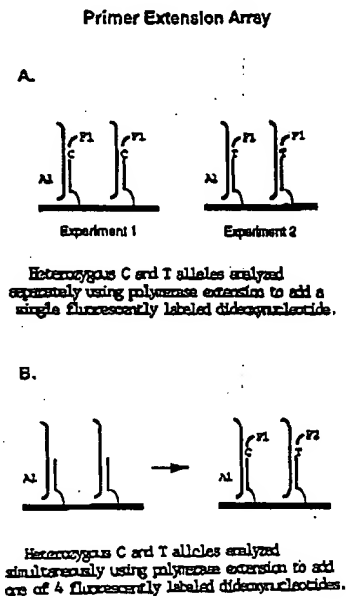


Fig. 4. DNA array using primer extension to detect putative mutation.

In this approach, the target sequence is hybridized to an array where each address represents a primer with the 3' end adjacent to the base which will be interrogated. When the reactions are performed separately on four arrays, DNA polymerase extension adds a single fluorescently or radioactively labeled dideoxynucleotide which is the complement of the hybridized sequence. When using four fluorescently-labeled ddNTPs, the reaction may be performed on a single array (Figure 4B). Examples using radioactive dNTP and fluorescently labeled ddNTPs is shown in Figure 4 of a review by Ed Southern, a world renowned expert on hybridization who developed the well known "Southern hybridization" technique (Southern, "DNA Chips: Analysing Sequence by Hybridization to Oligonucleotides on a Large Scale." TIG, 12:110-115 (1996) ("Southern") (attached hereto as Exhibit 6)). While the correct sequence can be determined, inappropriate base extensions are also observed and background noise is very high. Further, solid-phase primer extension also generates false

positive signals from mononucleotide repeat sequences, template-dependent errors, and template-independent errors. Therefore, the primer extension technique will not be able to unequivocally distinguish the presence of cancer mutations in primary tumors (where the mutation may represent 15% compared to 85% wild-type sequence), nor in clinical samples where the mutation represents only 1% or less of the target sequence.

15. Southern describes polymerase extension of oligonucleotide addresses, then he writes:

> Ligases can be used in a related fashion. In this case, one oligonucleotide is tethered, the other in solution; either one or both oligonucleotides can be used to provide sequence discrimination. The use of enzymes in a solid phase format enables introduction of label through the use of tagged triphosphate precursors or oligonucleotides, thus removing the need to label the target. This can be a great advantage when multiple targets are to be analyzed. Thus many of the methods that have been used to detect mutations using polymerases and ligases can be adapted for use with arrays. The indications are that discrimination is, indeed, greater than with hybridization alone.

Id. at 115. Thus, Southern teaches the use of solid-phase ligation to covalently attach a second oligonucleotide to a first oligonucleotide which is already tethered to the solid surface. This would appear to be a preferred method for combining ligase based detection of single nucleotide differences with an array since use of discriminating oligonucleotides as the actual addresses and solid-phase ligation on an array would reduce the potential problem of primer overlap in a solution-based multiplex ligase detection reaction. Thus, Southern teaches away from the use of solution phase ligase-based detection with an addressable DNA array.

16. U.S. Patent No. 5,695,934 to Brenner ("Brenner") discloses a procedure for massive parallel sequencing of different fragments of DNA. This technique involves (1) cloning fragments of DNA and sorting them onto beads or a two dimensional array; (2) sequencing the fragments by repeatedly labeling the fragments, identifying what base the label corresponds to, and removing one nucleotide from the fragment; and (3) assembling the fragments into a completed contig. To distinguish Brenner from the present invention, step (1) will be described in more detail. This involves a combinatorial approach

to synthesizing tags on beads where each bead has identical tags which differ from the tags on other beads. Both addresses and their complements are synthesized, and a portion thereof are cleaved from the beads and hybridized with each other to make the tags double stranded. The tag complements have minimally cross-hybridizing properties. The oligonucleotide tags are ligated to fragments of polynucleotides to be sequenced, placed into a vector, and cloned into *E. coli*. The *E. coli* is then permitted to grow so that numerous copies of this identical bacteria with copies of the oligonucleotide tag and the polynucleotide fragment are produced. The joined oligonucleotide tag and polynucleotide fragment from each colony is excised, and a biotin or *Fok*I restriction site is then placed at the end of this fragment which does not have the tagged oligonucleotide to facilitate the sequencing procedure of step (2). The tagged oligonucleotide is then made single stranded by digestion with a polymerase enzyme having 3' to 5' exonuclease activity in the presence of a single nucleotide triphosphate. In order for this procedure to work, the tag sequence must lack at least one of the 4 natural nucleotides such that polymerase degradation renders the entire tag sequence single stranded. The resulting product is contacted with beads to which tag complement is attached under conditions effective to permit hybridization of the tagged oligonucleotide to the complement tag. Ligase or some other chemical means is used to cross-link permanently the polynucleotide fragment to the bead. The product is then passed over a steptavidin coated slide and captured. The captured polynucleotides can then be sequenced. Thus, Brenner has nothing to do with a solution phase, ligase detection reaction based procedure, which utilizes a resusable addressable array, for detecting single nucleotide base differences.

17. Brenner's tag complements are also very different from those of the present invention. Brenner requires that double stranded tags within cloned DNA contain only 3 out of the 4 natural bases, so that they may be rendered single stranded by the aid of a polymerase enzyme. The addressable portions of LDR primers in our invention are readily distinguished from the Brenner tags, in that our addressable portions contain all four natural bases, and further do not require any enzyme or additional step to render them single stranded. Examples of Brenner's minimally cross-hybridizing set of subunits are provided in Tables I and II in column 7. Although table legends are missing, it is clear that the sets in Table II are read as columns, i.e., the first column has two sets, the first set containing 8 tetramer members, all lacking in "G", the second set containing 9 members, all lacking in "T". Columns 2 to 6 all have sets containing 18 tetramer members, all lacking "T". Let us

consider the set described in row 3, and number them 1-18 in order of appearance. All minimally cross-hybridizing 4mers in this set lack a T, and thus fall under the definitions described above for a preferred embodiment whose subunits are made up of three of the four natural nucleotides. This set differs substantially from the set of 36 tetramers in our application, which exhibit the following properties: (i) each tetramer differs from the others by at least two nucleotides; (ii) no two tetramers are complementary to each other; and (iii) palindromic tetramers have been eliminated. Note that Brenner tetramers #9 (GGCC) and #17 (GCGC) are palindromic. For Brenner, Table II, column 3, note that tetramers #1 (AAAC) and #13 (CAAC) differ by only one nucleotide. Likewise, tetramers #2 (ACCA) and #11 (ACAA) differ by only one nucleotide. Likewise, tetramers #3 (AGGG) and #10 (AAGG) differ by only one nucleotide. Likewise, tetramers #4 (CACG) and #14 (CCCG) differ by only one nucleotide. Likewise, tetraters #5 (CCGC) and #17 (GCGC) differ by only one nucleotide. Likewise, tetramers #6 (CGAA) and #15 (CGGA) differ by only one nucleotide. Likewise, tetramers #7 (GAGA) and #16 (GACA) differ by only one nucleotide. Likewise, tetramers #9 (GGCC) and #12 (AGCC) differ only by one nucleotide. Likewise, tetramers #8 (GCAG) and #18 (GGAG) differ by only one nucleotide. Numerous other examples of tetramers which differ from other tetramers by only one nucleotide in that set are also evident in columns 2, 4, 5, and 6. Thus, numerous tetramers in Brenner's minimally cross-hybridizing tetramer sets in columns 2 to 6 (as well as every tetramer in column 3) contradicts the teachings of our invention. Thus, Brenner most certainly teaches away from the design of addresses as described in the present invention.

18. The above art on DNA arrays does not teach how to detect single nucleotide mutations which arise spontaneously in cancer genes. In addition, my prior work (see ¶ 5) did not describe how to detect multiple adjacent and distant loci in multiple genes using PCR/LDR without an array.

19. In the present invention, we have developed a multiplex polymerase chain reaction / ligase detection reaction (PCR/LDR) method which identifies all 19 possible single-base mutations in K-*ras* codons 12, 13, and 61, with a sensitivity of 1 in 500 wild-type sequences. The K-*ras* gene presents the dual challenge of distinguishing multiple mutations in the neighboring codons 12 and 13 from wild-type sequence, as well as the closely homologous H-*ras* and, N-*ras* genes. Probes specific for each point mutation can interfere with one another during hybridization or polymerase extension. Thus, while allele-specific

PCR can detect individual point mutations, it is not effective for multiplex detection of all possible K-*ras* mutations in a homogeneous assay.

20. To approach the above challenges, a primary gene-specific PCR coupled to a multiplex LDR assay was developed. Independent K-*ras* specific PCR primers were designed for Exon 1, containing codons 12 and 13, and for Exon 2, containing codon 61. Twenty-six LDR probes were designed and synthesized for each of 19 possible point mutations in codons 12, 13, and 61. The discriminating oligonucleotides were fluorescently-labeled and contained the mutant base at their 3' end.

21. In a standard LDR reaction, an excess of LDR probe over template assures efficient hybridization and ligation during each cycle. However, when combining LDR probes to test all six possible single-base mutations at K-*ras* codon 12, the six upstream discriminating and two downstream common probes as well as the complementary PCR strand all compete for hybridization to the same region of template DNA. A balance must be sought between the ratio of total LDR probe and target DNA, such that sufficient probe is added to generate a robust LDR signal, but not so much total probe as to squelch (i.e. significantly reduce) the signal generated from the correct probe. Optimal results were obtained using the following conditions: the total amount of PCR amplified product (2,000 fmoles total product in 20 µl LDR reaction), fluorescently-labeled discriminating probes (500 fmoles each), common probes (from 500 to 1,500 fmoles each), and *Tth* DNA ligase (100 fmoles). For the eight probe set, the limit of positive identification of a G12V mutation was 1 cancer mutation in 1,000 wild-type templates with a signal-to-noise ratio greater than 3:1. Thus, probe interference during hybridization reduced the limit of positive detection at least four-fold, and reduced the total LDR product yield about five-fold in the six mutation multiplex reaction as compared to the simpler two mutation detection reaction. Until these experiments were performed, it was not known if 8 overlapping probes would work in an LDR reaction and provide sufficient signal above noise to detect the presence of one mutation in 1,000 wild-type templates.

22. In the completely assembled multiplexed LDR reaction, a total of 26 probes were used to test all 19 possible single-base mutations at K-*ras* codons 12, 13, and 61. For the 26 probe set, the presence of twelve overlapping discriminating probes and four overlapping common probes which hybridize to codons 12 and 13 predictably reduces total signal, nevertheless, the limit of identification was one G12V mutation in 500 wild-type

templates, with a signal-to-noise ratio of 3:1. Until these experiments were performed, it was not known if 16 overlapping probes would work in an LDR reaction and provide sufficient signal above noise to detect the presence of one mutation in 500 wild-type templates.

23. In a blinded study, 144 paraffin-embedded archival colon carcinomas were microdissected, and K-*ras* mutations determined by both dideoxy-sequencing and multiplex PCR/LDR. Results were concordant for 134 samples. The ten discordant samples were re-evaluated using higher sensitivity uniplex PCR/LDR, and the original multiplex PCR/LDR result was confirmed in nine of these ten cases. The last case was an exceedingly rare double mutation in the same codon and hence would require a different LDR probe to detect via PCR/LDR. Multiplex PCR/LDR was able to identify mutations in solid tumors or paraffin-embedded tissues containing a majority of wild-type stromal cells, with or without microdissection. Thus, multiplex PCR/LDR succeeds where dideoxy sequencing or a hybridization array, as described in Figures 3A-B, would fail.

24. Several PCR-based techniques can detect single-base mutations present in a minority population of human tumor cells: allele-specific PCR (AS-PCR, which includes MASA, PASA, MS-PCR, MAMA, or ARMS), primer mediated RFLP, or electrophoretic mobility-based methods. However, these amplification methods are susceptible to false-positive signals due to mis-extension of the mutant-specific primer on wild-type target. Although mis-extension is inefficient, the majority of primer is hybridized to wild-type DNA and is a substrate for this unwanted reaction. Several groups have emphasized the need to carefully optimize primer design and reaction conditions under which each AS-PCR reaction is performed, since these conditions greatly determine the degree of discrimination which can be attained between mutant and wild-type. This problem is compounded when a primary PCR amplification step is used, as polymerase errors in the initial amplification provide false templates which can be subsequently selected and amplified into false-positive signal. Since AS-PCR and restriction digest techniques cannot accurately quantify the amount of minority template, false positive signals from polymerase errors or minute contamination cannot be easily distinguished from true positive signals. Use of mis-matched primers to create restriction sites reduces the fidelity of *Taq* polymerase extension of the initial base, generating additional false positive signals. In addition, neither AS-PCR nor restriction digests nor electrophoretic separation can precisely identify the full spectrum of mutations at a given gene locus. Multiplexing of AS-PCR is limited by primer

interference, which reduces the yield of correct product and increases the likelihood of mis-extension errors. Finally, neither AS-PCR nor restriction digests nor electrophoretic separation are compatible with an array-based detection format.

25. PCR/LDR overcomes many of the above limitations by separating the amplification (i.e. PCR) and mutation discrimination (i.e. LDR) steps. A distinguishing feature of PCR/LDR is that mis-ligations do not undergo subsequent amplification, therefore reducing the chance of false positive reactions. Any low-level polymerase errors remain unselected, and thus contribute only a minimum of background noise. Since PCR/LDR also tests for the presence of any base change (i.e. in K-*ras* codon 12), a profile of background noise due to mis-ligations is generated, which may serve as internal markers allowing true-positive signal to be distinguished.

26. When compared with direct DNA sequencing, PCR/LDR successfully identified the full spectrum of point mutations in codons 12 and 13. Among 96 tumors identified as wild-type by direct sequencing, K-*ras* mutations were identified in eight of these tumors using multiplex PCR/LDR, and in all cases the presence of the mutation was confirmed by repeat PCR and high sensitivity LDR using a single probe pair. In only one tumor were two K-*ras* point mutations identified, suggesting that double mutations are a rare event when, as in this study, carcinomas are sampled at a single locus by microdissection. PCR/LDR was also successful in analyzing archival tumors from whole paraffin sections without microdissection as well as from frozen tumor tissue. Taken together, the data support the conclusion that PCR/LDR has the decided advantages of (i) allowing large-scale multiplexing, (ii) providing quantitative detection of mutations in a high background of normal sequences (iii) allowing detection of closely-clustered mutations, (iv) being amenable to automation.

27. We have pursued a strategy for high-throughput mutation detection which is substantially different from the DNA array designs discussed above. Combined with solution-based polymerase chain reaction / ligase detection assay (PCR/LDR), this array allows for accurate quantification of single base mutations, whether inherited and present as 50% of the sequence for that gene, or sporadic and present at 1% or less of wild-type sequence. For high throughput detection of specific multiplexed LDR products, we attach unique addressable portions to our LDR probes to guide each LDR product to a designated address on a DNA array. The specificity of our reaction is determined by a thermostable

ligase which allows detection of (i) dozens to hundreds of polymorphisms in a single-tube multiplex format, (ii) small insertions and deletions in repeat sequences, and (iii) low level polymorphisms in a background of normal DNA. By uncoupling polymorphism identification from hybridization, each step may be optimized independently, thus allowing for quantitative assessment of allele imbalance even in the presence of stromal cell contamination. This approach has the potential to identify rapidly low abundance mutations in multiple codons in multiple genes, as well as multiple gene deletions and amplifications associated with tumor progression.

28. Figures 5A-B below present two alternative dual label strategies to quantify LDR signal using addressable DNA arrays.

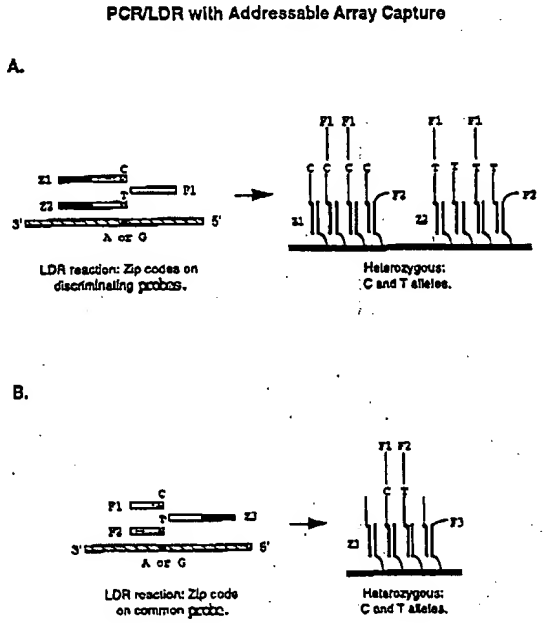


Fig. 5. Addressable DNA array using PCR/LDR to detect putative mutation.

In Figure 5A, the common LDR probe for both alleles contains a fluorescent label (F1) and the discriminating probe for each allele contains a unique addressable sequence. Following hybridization of the LDR reaction mixture to an array composed of fluorescently-labeled (F2)

zip-codes, the ratio of F1/F2 for each address can be used to determine relative percent mutation or allelic imbalance. In Figure 5B, the common probe for both alleles contains an addressable sequence and the discriminating probe for each allele contains a unique fluorescent label, F1 or F2. Following LDR, the reaction mixture is hybridized to an array and the ratios of F1/F2 for each address can again be used to determine relative percent mutation or allelic imbalance. In addition, by adding a third label, F3, to the oligonucleotide address coupled to the surface it is possible to quantify each allele separately.

29. Figures 6A-B below show the scheme for PCR/LDR detection of mutations in codons 12 and 13 of K-*ras*. using an addressable array.

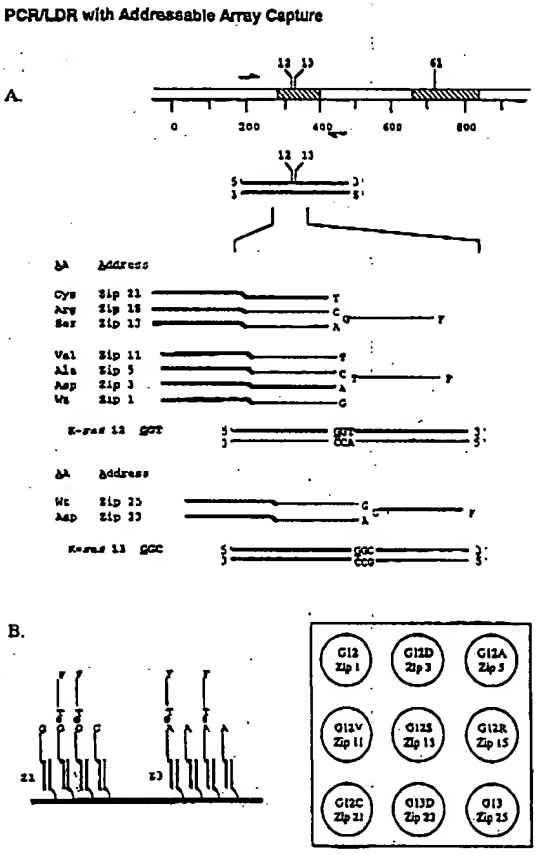
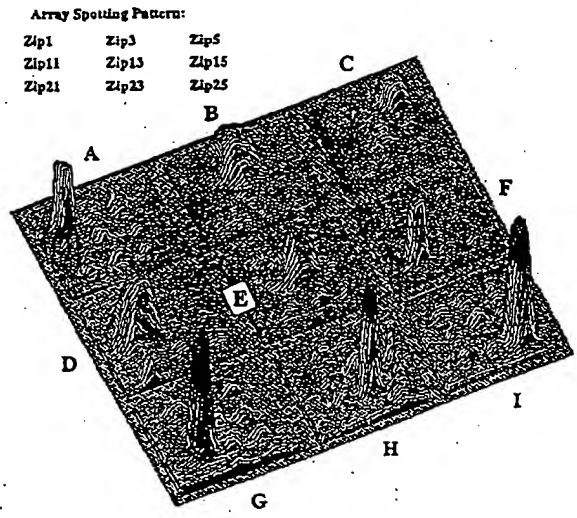


Fig. 6. LDR Probe sets for addressable DNA array detection of K-*ras* mutations.

Figure 6A shows a schematic representation of chromosomal DNA containing the K-*ras* gene. Exons are shaded and the position of hot-spot codons 12 and 13 are shown. Exon-

specific PCR primers were used to selectively amplify K-*ras* DNA flanking codons 12 and 13. Probes were designed for LDR detection of the seven most common mutations found in the K-*ras* gene in colon cancer. For example, codon 12, second position, (GGT) may mutate to GAT, GCT, or GTT. Allele-specific LDR probes contained a complement to an addressable sequence on the 5' end and the discriminating base on the 3' end. Common oligonucleotides were phosphorylated on the 5' end and contained a fluorescent label (F) on the 3' end. Different mutations were distinguished by separating the products on an addressable DNA array (Figure 6B). LDR probes used for detecting mutations at codons 12 and 13 compete for hybridization to discrete addresses on the array. Only probes containing the correct complement to a given address remain bound, giving rise to a fluorescent signal. Surface plots showing array detection of K-*ras* mutations from cell line DNA are shown below in Figure 7.

Surface Plots of Arrays Hybridized with 10 Probe Set LDR Reactions



Array	Mutation	Codon	Mutant/Wild Type Ratio
A	Wild Type	12	0
B	Aspartic Acid	12	1.0
C	Alanine	12	0.65
D	Valine	12	4.1
E	Serine	12	0.9
F	Cysteine	12	0.7
G	Cysteine	12	0.7
H	Aspartic Acid	13	1.1
I	Wild type	13	0

Fig. 7. Addressable DNA array using PCR/LDR to detect K-*ras* mutations.

Each array was hybridized with an individual LDR reaction which contained cell line DNA with a known mutation and a 10 probe set. LDR reactions detecting wild type DNA (arrays A and I) used only a two probe set specific for normal DNA (wild-type) at either codon 12 (Array A) or 13 (Array I). The position of the hybridization signal was correct for all nine arrays. Variance in shape is a consequence of manual spotting.

30. To demonstrate DNA array capture sensitivity, mixtures of an excess of unlabeled to labeled probes were tested. This is equivalent to an LDR reaction where an excess of unligated probes compete with the labeled LDR products for hybridization to the array. DNA arrays were hybridized in quadruplicate with from 100 amoles to 30 fmol of a synthetic 70-mer LDR product mixed with a full set of K-*ras* LDR probes (combined total of 9 pmol of discriminating and common probes) under standard conditions. The arrays were analyzed on two different instruments: a Molecular Dynamics FluorImager 595 and an Olympus AX70 epifluorescence microscope equipped with a Princeton Instruments TE/CCD-512 TKBM1 camera. We can achieve a signal-to-noise ratio of greater than 3:1 starting with a minimum of 3 fmol synthetic 70-mer LDR product labeled-probe within 4,500 fmol Fam label and 4,500 fmol address containing LDR probes in the hybridization solution. See Figures 8A-B below.

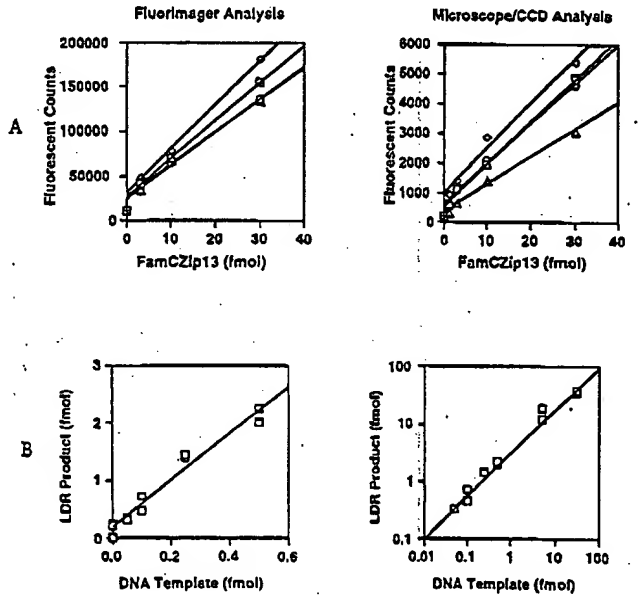


Fig. 8. A: Quantification of minority fluorescently-labeled oligonucleotide containing an addressable portion captured by a universal addressable array using two different detection instruments. B: Quantitative detection of G12V mutation of the K-*ras* gene using two LDR probes in the presence of 10 microgram salmon sperm DNA.

R203863.1

Initial results from the microscope/CCD instrumentation indicate we can maintain a 3:1 signal-to-noise ratio starting with 1 fmol labeled product, and with optimization may be able to extend the limit to 100 amol. For both instruments, a linear relationship is observed between labeled synthetic 70-mer LDR product added and fluorescent counts captured. This result was unanticipated since one might predict that the majority unlabeled LDR probes containing addressable portions would compete with the minority labeled LDR products containing addressable portions to the extent that low amounts of LDR product might not be detectable.

31. Each array was plotted individually, and variation in fluorescent signal between arrays may reflect variation in amount of oligonucleotide coupled due to manual spotting. Rehybridization of the same probe concentration to the same array is reproducible to +/- 5%, with capture efficiency from 20 to 50%. Since the total of both labeled and unlabeled probe containing addressable portions which complements a given address remains unchanged (at 500 fmol) from LDR reaction to LDR reaction, this result demonstrates the ability to quantify the relative amount of LDR product using addressable array detection. Since the relationship between starting template and LDR product retains linearity over 2 orders of magnitude with a similar limit of sensitivity at about 100 amols (see Figure 8B), combining PCR/LDR allele discrimination with array-based detection provides quantifiable results.

32. The addressable DNA array may also be used to quantify LDR product when the initial sample contains a minority of mutant DNA in an excess of wild-type sequence. A dilution series was prepared containing wild-type K-*ras* DNA and SW620 cell line DNA (G12V) in ratios of 20:1 to 500:1. LDR was carried out on 2,000 fmol total DNA using 2,000 fmol G12V discriminating probe and 2,000 fmol 3'-Texas Red labeled common probe. Arrays spotted with aid of a robot were hybridized in quadruplicate. These arrays consisted of 200 micron spots with 4 spots per address. The arrays were analyzed on an Olympus AX70 epifluorescence microscope equipped with a Princeton Instruments TE/CCD-512 TKBM1 camera. This experiment demonstrated quantitative detection of G12V LDR product on an addressable array at ratios as low as 1 mutant target in 200 wild-type sequences. See Figure 9 below.

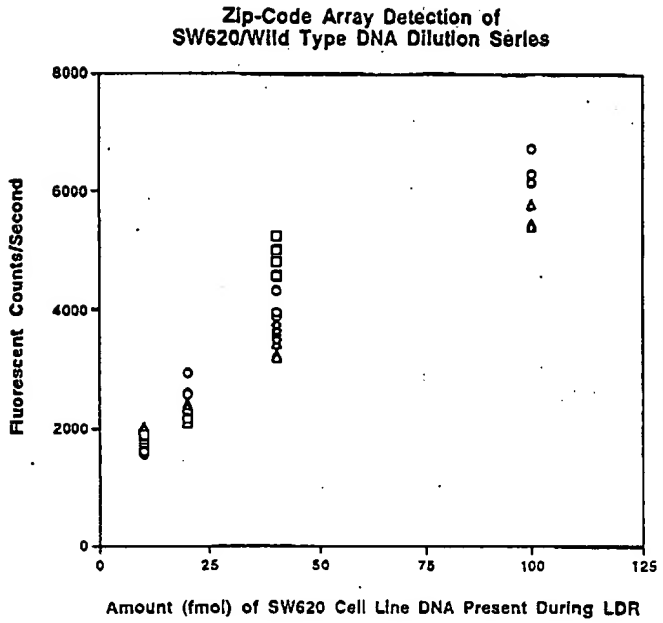


Fig. 9. Detection of minority K-*ras* mutant DNA in a majority of wild-type DNA using PCR/LDR with addressable DNA array capture. DNA from cell line SW620, containing the G12V mutation, and DNA from normal lymphocytes was PCR amplified in Exon 1 of the K-*ras* gene. Samples containing 10 fmol, 20 fmol, 40 fmol and 100 fmol of SW620 amplified fragment and 2,000 fmol of PCR amplified wild-type fragment in each were mixed and presence of mutation determined using LDR probes specific for the G12V mutation. Symbols indicate results from quadruplicate hybridization's on four independent DNA arrays.

Signal could be distinguished at the 1 in 500 ratio; however, dust particles in the DNA array matrix interfered with quantification. This technical difficulty is surmountable using standard "clean room" procedures. Thus, we demonstrate combining PCR/LDR with addressable array capture detects a low abundance cancer mutation present at less than 1% of wild-type DNA.

33. Fodor '305, Pease, and Cronin teach one approach to synthesizing arrays with oligonucleotides and a target-hybridization based approach to attempt to discriminate single nucleotide differences. The procedure for constructing these arrays simply combines known photolithographic techniques (well understood for making computer chips) with standard photoprotected monomer nucleoside chemistry. The work of Pease

claims a "surface chemical coupling yields efficiency ranging between 85% and 98%" per cycle. Id. at 5024. In Fodor '305, the only mention of coupling yields is on column 17 line 61 where it is stated "The coupling yield per cycle in these experiments is typically between 85% and 95%". Such an error rate would imply that the best and worst synthesis for 15 mer addresses (used by Cronin) are $(95\%)^{15} = 46\%$ correct sequence and $(85\%)^{15} = 8.7\%$ correct sequence. While failure sequences or deletions in the addresses may increase noise levels, there is still sufficient address for capture of PCR amplified probe. That is because a typical PCR reaction starts with 50 picomoles of each primer and yields several picomoles of product, sufficient for capture on an array, even if the address is mostly incorrect. In contrast, in the present invention, the preferred addresses are 24-mer in length. Were the procedure of Fodor '305 used to synthesize 24-mer addresses, the best and worst synthesis are $(95\%)^{24} = 29\%$ correct sequence and $(85\%)^{24} = 2\%$ correct sequence. A typical LDR reaction starts with 500 femtomoles of each discriminating LDR probe, or about 100-fold less LDR primer than a PCR reaction. This is one reason why LDR reactions may be multiplexed to contain hundreds of LDR probe sets, each individual probe is at such low concentrations. However, the amount of LDR product formed in a typical reaction (300 attomoles to 50 fmoles, see Figure 8B) is only a fraction of the starting LDR product. In our array-based format, the LDR product competes with unreacted probes for hybridization to the same address. Thus, an LDR reaction starts with 100-fold less probe than a PCR reaction, only a fraction of this LDR probe is converted to LDR product, and the LDR product is further competed in hybridizing to the array -- i.e. only a small amount of product hybridizes. The poor yields for making arrays as taught in the above references would be expected to provide insufficient addresses of the correct sequence on the array for efficient capture of our labeled LDR products containing addressable sequence portions in the presence of competing non-labeled LDR probes containing the same addressable sequences. Therefore, the art of Fodor '305, Pease, and Cronin teach away from the present invention.

34. Zaun '839 discloses an apparatus and method for amplifying and detecting target nucleic acids. This procedure involves amplifying with a thermal cycling device and then detecting reaction products on a support having one or more reaction sites. Amplification can be carried out using PCR or LCR procedures. To capture amplification products, the detection system is provided with a support having a plurality of capture sites to

immobilize such products on the support. Zaun discloses capturing amplification products with antibody-antigen binding, chemical bonding, or complementation of polynucleotides.

35. As noted above, Zaun, at most, relates to use of LCR as an amplification procedure which is distinguishable from LDR. Moreover, although Zaun refers to a number of LCR procedures (see col. 1, lines 50-62), the LCR procedure being practiced in that reference (see col. 39, lines 1 to 11) is very different from that described in my papers in ¶ 5. More particularly, the LCR procedure utilized by Zaun references EP-A-439,182 to Backman, et al., (which is already of record) which is known as gap-LCR. The difference between gap-LCR and the LCR procedure set forth in my papers identified in ¶ 5 (compare Figures 1 and 4). In brief, my LCR procedure involves use of oligonucleotide probe sets which hybridize to a target nucleic acid in abutting relationship, and, if there is perfect complementarity at their junction, these oligonucleotides can be joined with ligase. By contrast, in gap-LCR, the oligonucleotide probe sets do not hybridize to a target nucleic acid in abutting relation and, therefore, are not potentially suitable for immediate ligation. Instead, there is a gap which must be closed using polymerase before any ligation can occur.

36. This difference between my LCR procedure and gap-LCR is more than a mere difference in methodology. This difference has a significant impact on the utility of each technique. For example, in detecting sickle cell anemia, the distinction between the normal β^A and sickle β^S genes is a single A → T transversion which leads to a change from a glutamic acid residue to a valine in the hemoglobin β chain. This single base difference is readily detected with my LCR procedure. Gap-LCR is able to distinguish the single base difference between *L. monocytogenes* and *L. innocua*. See Figure 4 of Wiedmann. However, when this procedure is applied to detection of sickle cell anemia, it is ineffective, because it forms the same ligation product when the β^A globin gene and β^S gene are present. Where the base differences between matched and mismatched targets are <u>not complementary</u>, gap-LCR can be used to detect single base differences. However, when the bases distinguishing matched and mismatched target are <u>complementary</u> (e.g., A for the β^A globin gene and T for the β^S globin gene), gap-LCR is ineffective.

37. The claims of the present application clearly refer to an LDR procedure to distinguish single nucleotide differences and require that the oligonucleotide probes be configured to hybridize "adjacent to one another" on a corresponding target nucleotide sequence. Since they do not involve LCR or filling a gap, these claims are readily

distinguishable from Zaun. Since Zaun cannot be used to distinguish all possible single-nucleotide differences, and Zaun never considers distinguishing any single-nucleotide differences, Zaun in no way suggests the present invention.

38. Cronin and Guo teach the use of PCR to amplify target DNA. However, a standard PCR product will not work well in hybridization. Both Cronin and Guo teach conversion of the PCR product to single strands, Cronin by asymmetric PCR and product fragmentation, Guo by streptavidin capture of biotinylated primer and removal of that strand. Guo states:

> Although it is simpler to prepare double-stranded PCR products than single-stranded, hybridization of the double-stranded molecule to the support will necessarily suffer from competition of the complementary strand with the support bound oligonucleotide. To evaluate this, a comparison of hybridization efficiency using both single-stranded and double-stranded 157 nt. PCR products were performed, with the results shown in Figure 3c. It is clear from these data that hybridization efficiency is much greater with the single-stranded than the double stranded product.

Id. at 5460. In the present invention, PCR products are used in the LDR reaction without fragmentation, without an asymmetric PCR step, and without a streptavidin capture of biotinylated primer step. Thus, Cronin and Guo teach away from the present invention.

39. Pease, Cronin, and Guo teach the use of a DNA address with a sequence complementary to the target with the discriminating base at or near the center. Cronin states:

> Each set of four probes is identical except for a single, central nucleotide position that corresponds to the target nucleotide to be identified. Under stringent hybridization conditions, a target nucleic acid molecule will hybridize best to its exact complement. The strongest hybridization signal will come from that particular probe in each set of four, permitting identification of the nucleotide at that position in the target.

Id. at 245. Extension of this teaching to detection of an LDR product on a DNA array suggests the appropriate address would contain the target sequence with the match or mismatched base at or near the center of the address. For example, an LDR product would

hybridize strongly to a perfectly complementary 20 mer address centered at the ligation junction. This would appear to be a preferred method for combining ligase based detection of single nucleotide differences with an array since only the LDR product would be of sufficient length to hybridize to an address while the initial LDR primers complement only the left half or right half of the address, and would not compete at the hybridization temperature. Further, this method would provide the advantage of discrimination of single nucleotide differences at both the ligation step and the array hybridization step. Thus, Pease, Cronin, and Guo teach away from the present invention.

40. Pease, Cronin, and Guo also do not teach the detection of low abundance mutations. While Pease demonstrates some selectivity in hybridization of a synthetic octamer to an array, the example chosen was a special case of hybridizing an octamer sequence of only "G" and "C" bases (GCGGCGGC) to an array at 15°C. Since it is well known that "G" and "C" bases in an oligonucleotide increase the oligonucleotide melting temperature (T_m), this example was biased in favor of the expected result. Even so, the closest false hybridization signal gave a signal-to-noise ratio of 5:1. Less biased examples are provided in Cronin using specialized DNA arrays designed with 40 addresses to determine presence or absence of a single point mutation. In the three examples provided: the perfect match signal was 92-100 and the mismatch signal was 22-33 for a signal-to-noise of about 1:3 to 1:4; the perfect match signal was 331-373 and the mismatch signal was 83-121 for a signal-to-noise of about 1:3 to 1:4; the perfect match signal was 123-150 and the mismatch signal was 30-41 for a signal-to- noise of about 1:3 to 1:4. While Guo claims signal-to-noise values for five mutations from 1:4 to 1:27, those mutations represent a select set which were chosen to have 5-7 GC content out of 15 nucleotides. These are the best cases for determining the presence of single nucleotide differences in samples where the mutation represents 50% of the total DNA present for that gene sequence. The signal-to-noise values of Pease, Cronin, and Guo would not be sufficient to detect mutations in primary tumors with stromal cell contamination, i.e. mutant DNA is present at 15% of wild-type DNA, or in other clinical samples where mutant DNA target is present at 1% of wild-type DNA. Thus, Pease, Cronin, and Guo do not suggest the present invention for detection of low abundance mutations.

41. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and

further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: 8/25/98

Dr. Francis Barany

Francis Barany